

The Potential of Administrative Data Linkage for Economic Policy in the UK: Experimental Results

**Presentation at the
ESCoE Conference 2019**

Russell Black, Silvia Lui, Philip Wales

Introduction

- A growing number of National Statistical Institutes and Central Banks are taking an interest in business registers, not as survey sampling frames, but as longitudinal records of business demography.
- Longitudinal databases that cover a universe of the businesses/employers on the business registers have been built. Examples:
 - The US Longitudinal Business Database by the US Census Bureau
 - The Canadian Longitudinal Employment Analysis Program by Statistics Canada
- These databases contain information on payroll, employment, industry characteristics, size and location, etc.

Introduction

- Some existing (non-UK) studies have shown the power of using these population datasets to improve our understanding of economic developments:
 - On the topics of: e.g. job creation and destruction, contribution to output and productivity growth, change in business dynamism, firm size distortion and productivity distribution
 - See, for example, Haltiwanger, Jarmin, and Miranda (2010), Haltiwanger, Jarmin, Kulick and Miranda (2015), Decker, Haltiwanger, Jarmin and Miranda (2016), Garicano, Lelarge and van Reenen (2016)

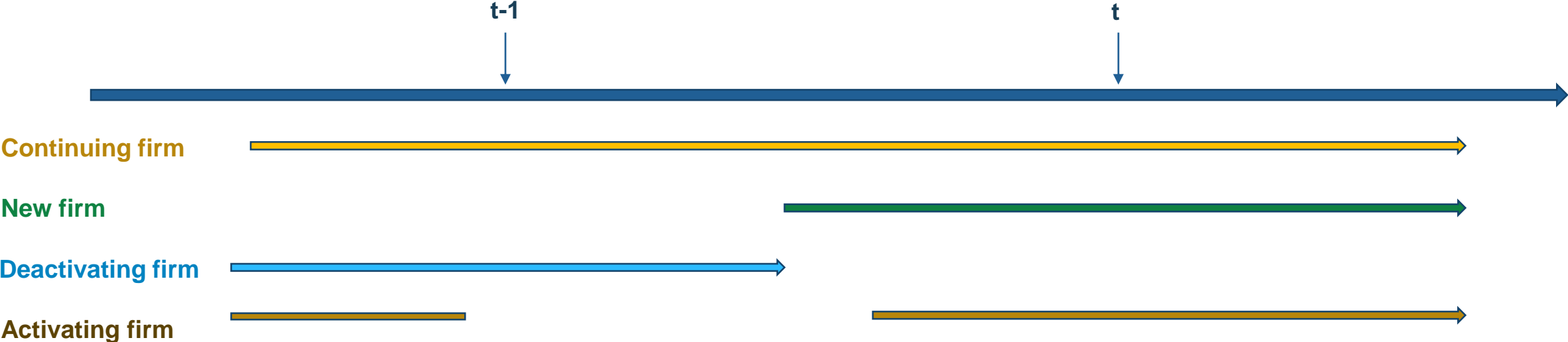
Introduction

- Existing studies using linked business datasets/database in UK:
 - On the topics of: e.g. firm dynamics; job creation, destruction and contribution.
 - Anyadike-Danes, Bonner and Hart (2011), Hart and Prashar (2019), Anyadike-Danes, Hart and Du (2013), Butcher and Bursnall (2018), Barnes and Haskel (2002)
- These studies usually use the Business Structure Database (BSD) or the Annual Respondents Database (ARD), or linking them to surveys to form longitudinal datasets.

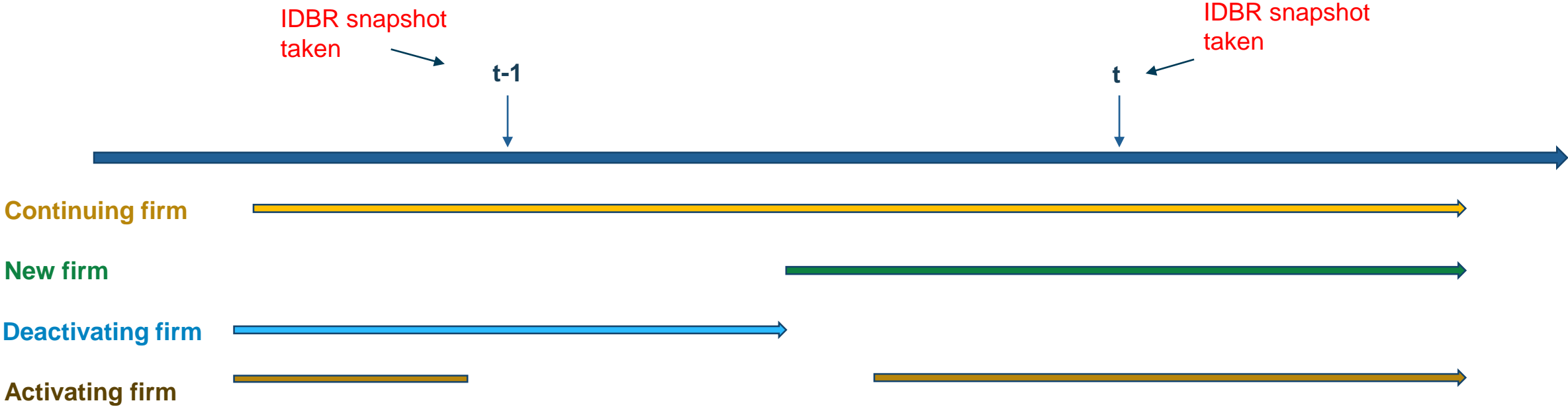
Characteristics of existing UK business databases

- Annual frequency
- Does not track businesses over time
 - The ABS is used for the ARD. In each year, ABS retains only half of the sample from the previous year
- Does not allow for linking data collected from multiple sources
 - Only allows linking data from survey source as administrative units of a firm is not included
- Does not truly capture continuity within a reference period due to the use of individual IDBR snapshot data
 - See example to illustrate the issue of using individual snapshot

Issue of considering individual snapshot to construct database



Issue of considering individual snapshot to construct database

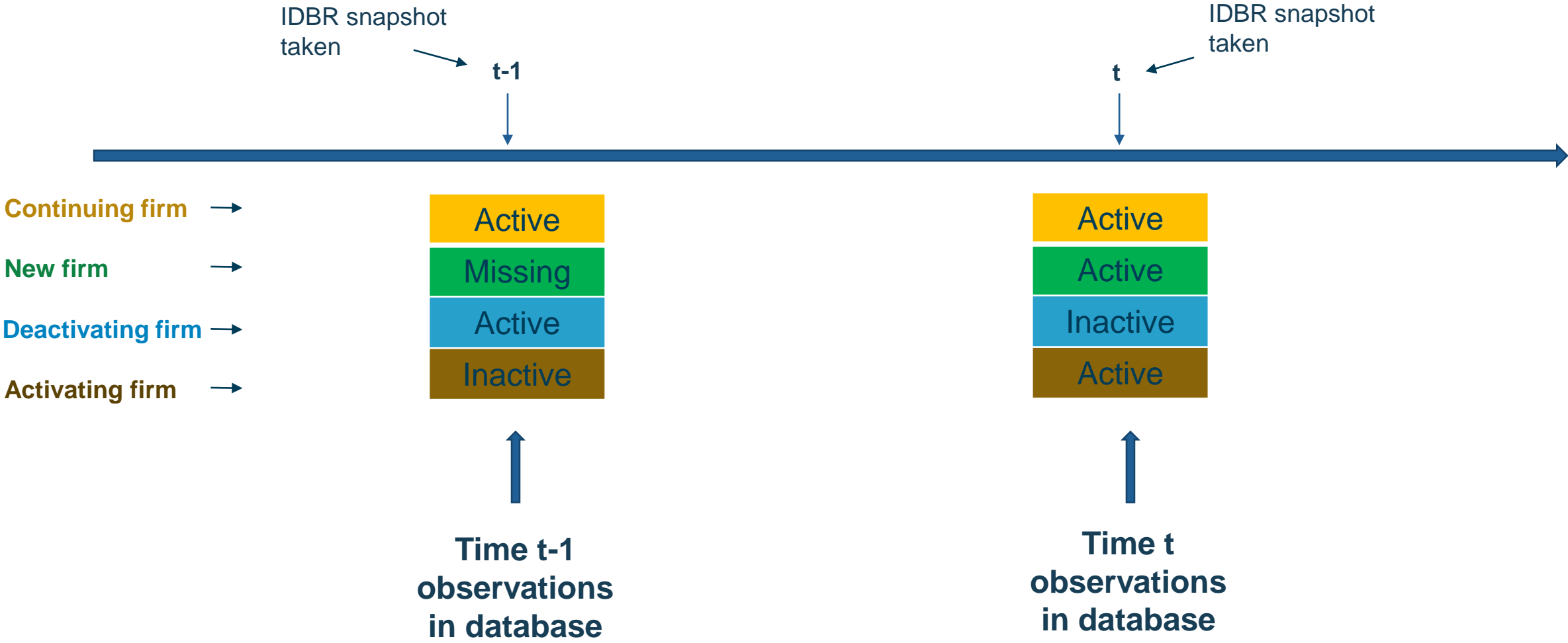


Issue of considering individual snapshot to construct database

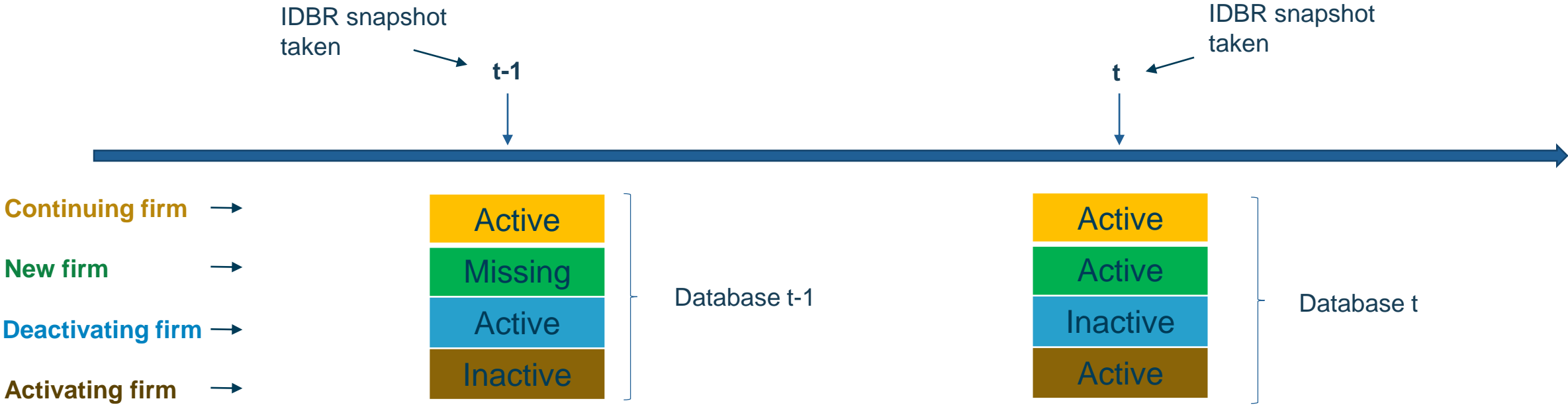


- Snapshot shows what is observed at the point in time when the “picture is taken”, but not what is observed over a reference period.

Issue of considering individual snapshot to construct database

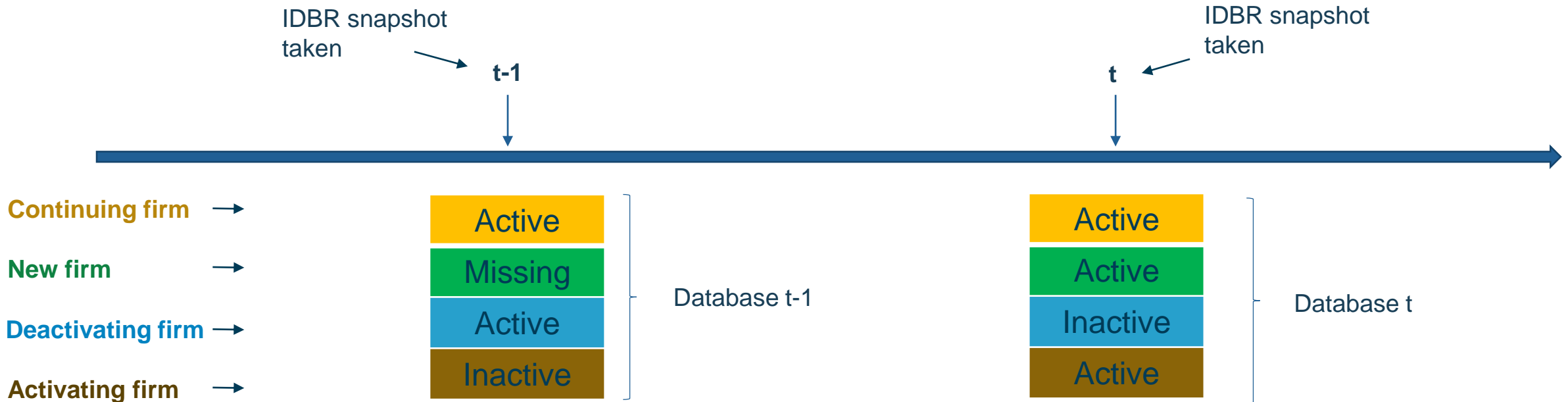


Issue of considering individual snapshot to construct database



- IDBR is survey sampling frame. Firms that are inactive (dead) on the register at the point in time when the sample is taken is not chosen.

Issue of considering individual snapshot to construct database



Example: Linking Survey C sampled between $t-1$ and t to the firms on the database. But no IDBR snapshot is taken between this two points in time to construct a cross-section of the database

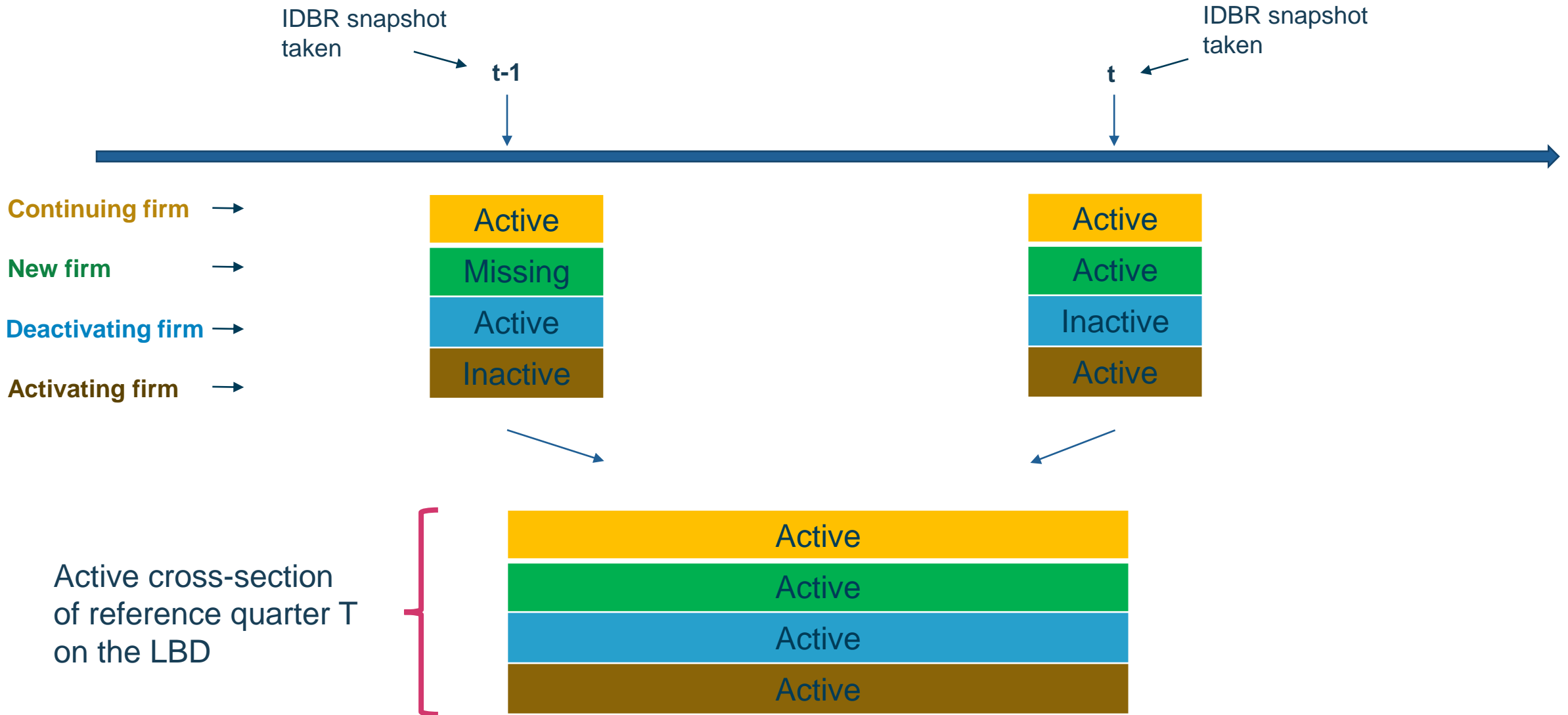
➤ Link to set of firms in the database in $t-1$ or t ?

The UK Longitudinal Business Database (UK LBD)

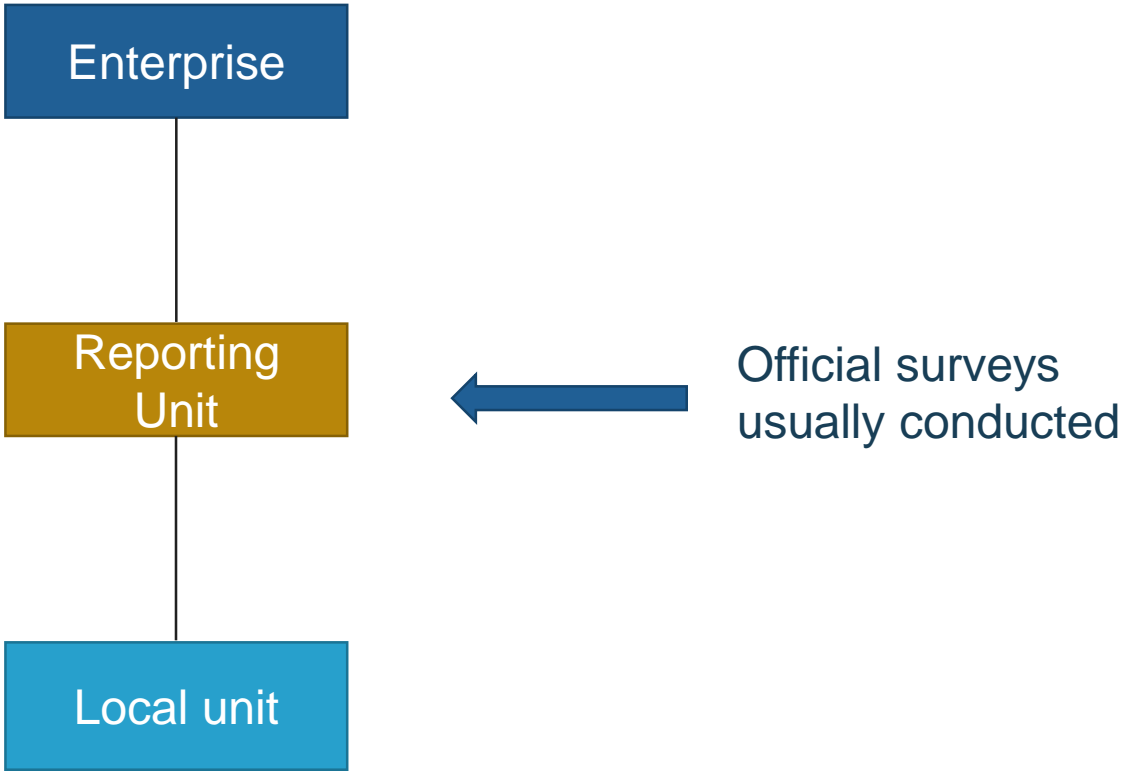
An ongoing project at the ONS aims to construct a quarterly LBD – an improvement to existing business databases:

- Near-population coverage with quarterly frequency
- Construct using consecutive IDBR snapshots
 - Capture continuity within a reference quarter
 - Capture within-quarter changes of industry and business structure
- Enable multi-level linking of data from both survey and administrative source
 - A single entry of a firm in a reference quarter displays the identity-links of all of its business units registered on the IDBR

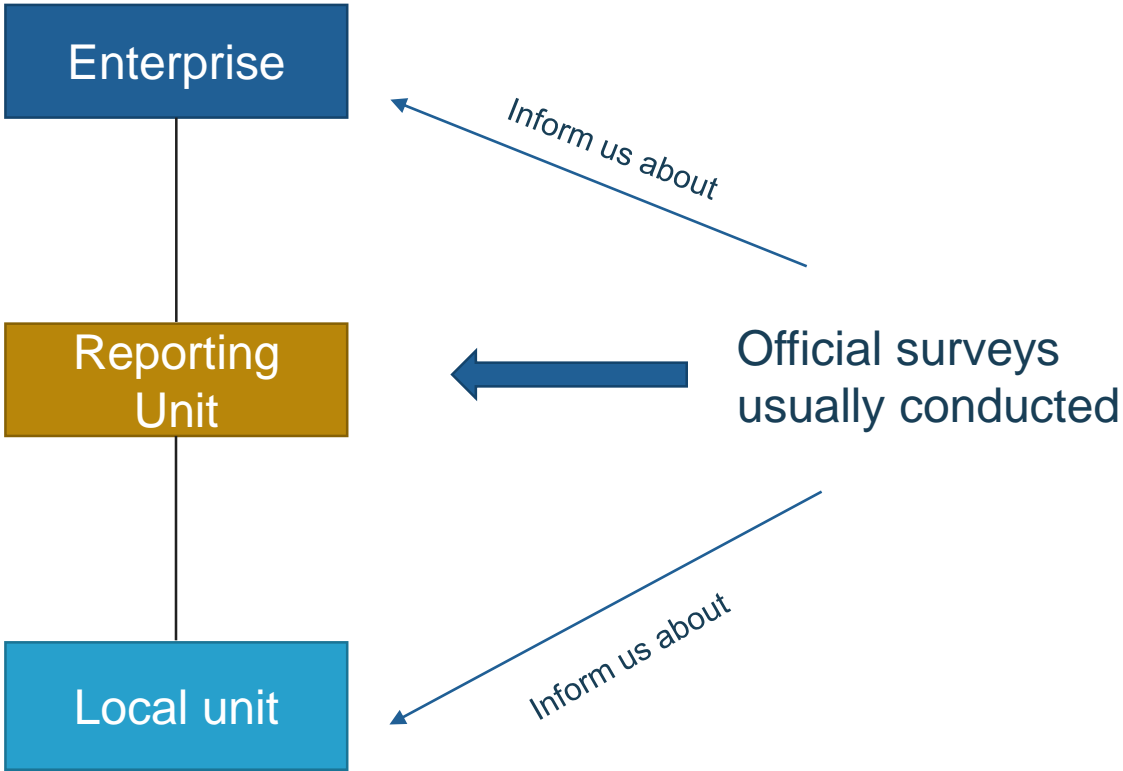
Each LBD cross-section is created using of consecutive snapshots



Multi-level data linking: a simple business example

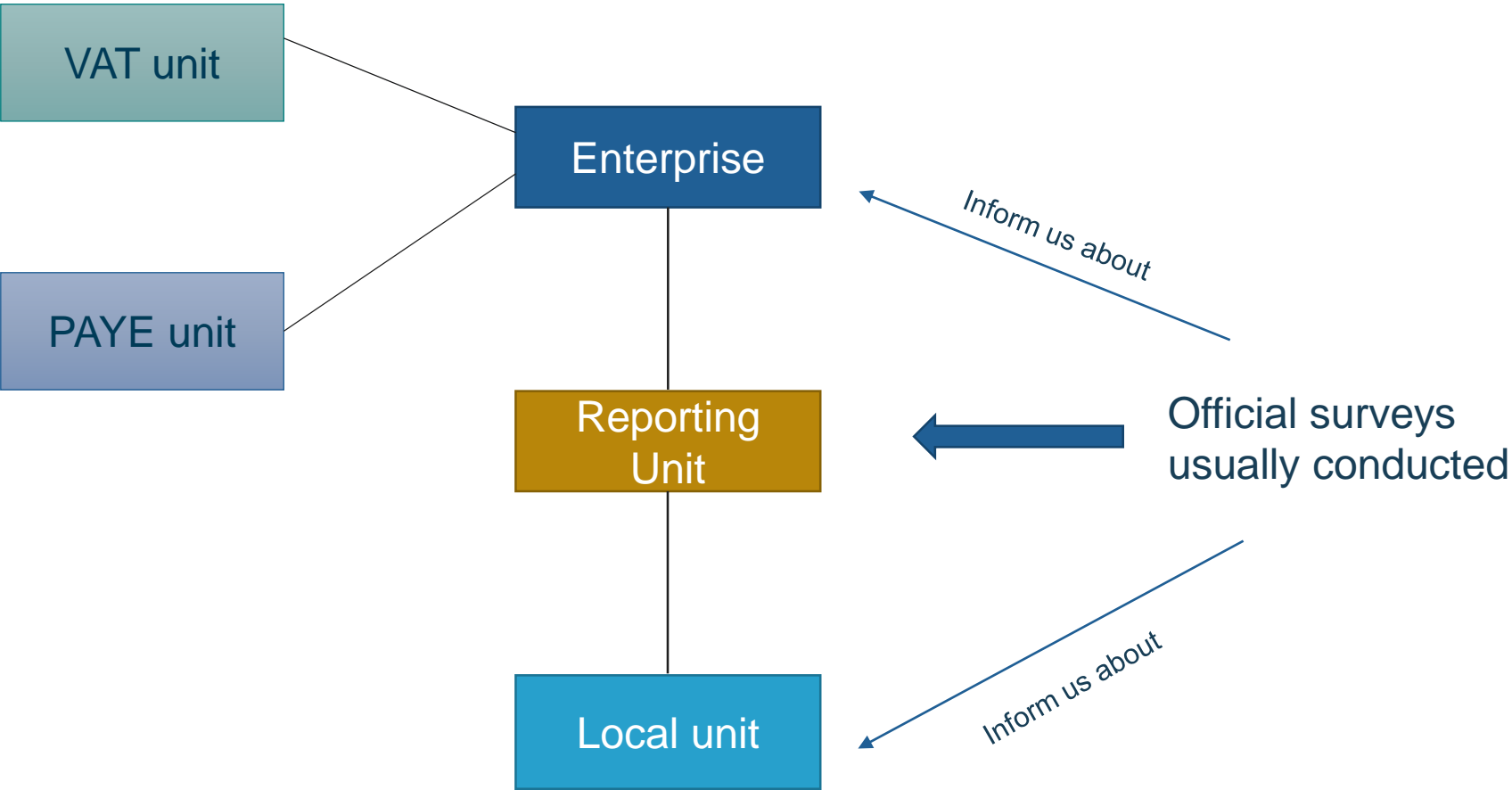


Multi-level data linking: a simple business example



Existing UK business databases consider only these units.

The LBD: A simple business example



The LBD also consider the two administrative units (VAT and PAYE) on the IDBR.

The UK LBD: A simple business example

Turnover and Trade data collected from administrative source



Employee data collected from administrative source



Inform us about



Official surveys usually conducted

Inform us about

The LBD has the potential to allow for linking with administrative data.

Analysis: Active Enterprise Population

- First attempt to examine business dynamism at the enterprise level using these experimental datasets from a cross-sectional point of view.
- These results are preliminary. They reflect what we have observed so far and we are still improving them.
 - We face challenges of using administrative data to produce this kind of analysis. For example, how these data are stored, how they should be used and the timing when the register is updated with these data, etc – We are still looking into how we can resolve these challenges in the future.
- Analysis is conducted using the quarterly populations of active enterprises.
- Active population in each reference quarter is constructed using two consecutive individual IDBR snapshots.
 - Individual IDBR snapshot is taken in a month of a calendar quarter
 - Hence, e.g. active population in reference quarter 2018Q3 (T) is constructed using IDBR snapshots taken in June (t-1) and September (t) 2018

Analysis

- Modifies the criteria adopted by the ONS's Business Demography to quarterly IDBR snapshots to identify a pool of active enterprises in a reference quarter.
- Compare enterprise-level turnover, employment, counts of live VAT and PAYE units in two consecutive snapshots to deduce signs of activity
 - An active enterprise is the one that has been active at some point during a reference quarter.
- Analysis uses employment data from PAYE source
 - Use employee counts (not individual-level data) on the IDBR collected from PAYE
 - When PAYE employment information is not available, we use enterprise employee counts on the IDBR that are collected from BRES

Transition

- An enterprise does not “die” but it can become active/inactive

	T-1	T	T+1
Activator in T	inactive	active	active
Continuer in T	active	active	active
Act-Deact in T	inactive	active	inactive
Closing in T, Deactivator in T+1	active	active	inactive
Deactivator in T, Closing or Act-Deact in T-1	active	inactive	inactive

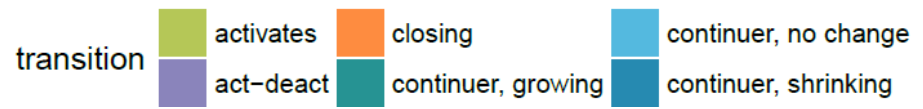
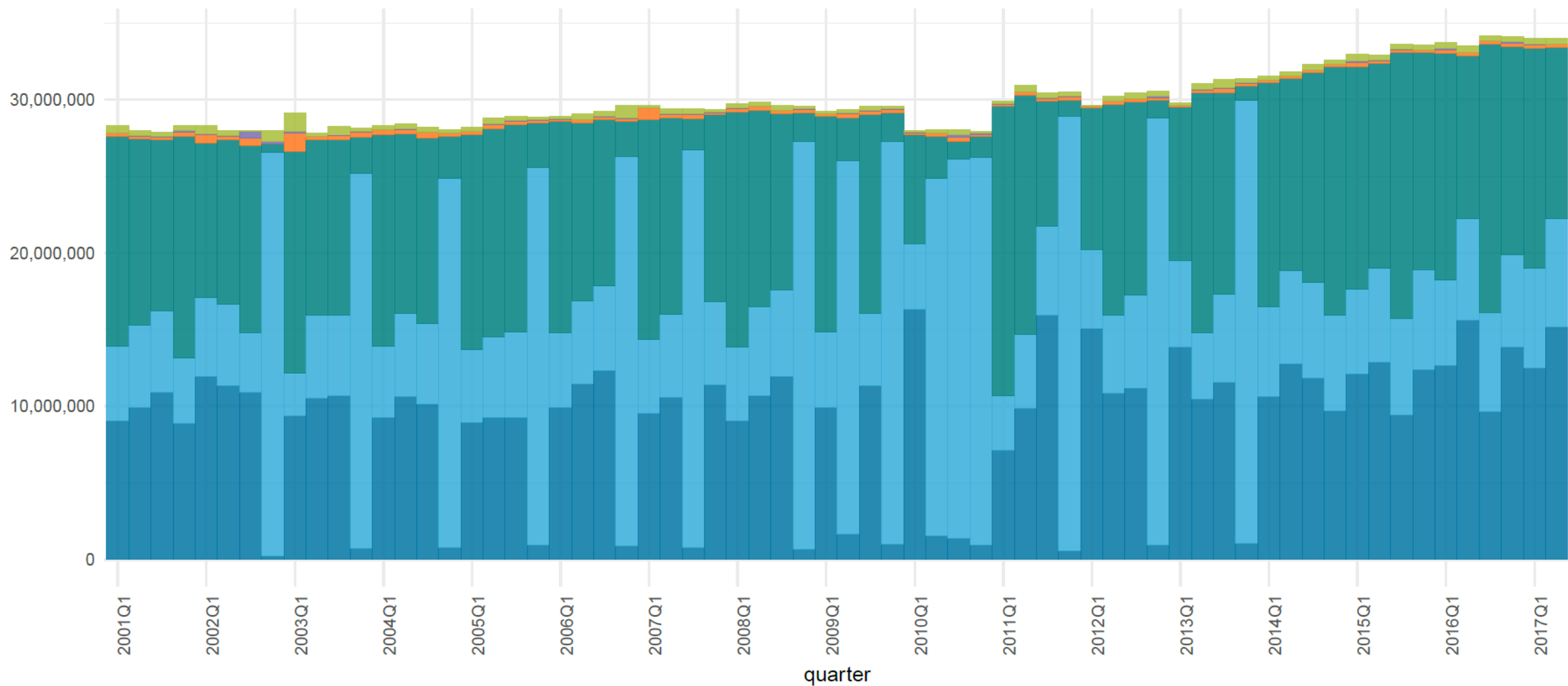
Look at employment in two ways

- Levels - of total employment in active enterprises, how many in...
 - Activators
 - Continuers
 - Closers
- Contributions (changes/first derivative) - change to total employment in active enterprises due to...
 - Activations
 - Deactivations
 - Continuers
 - (Closers)

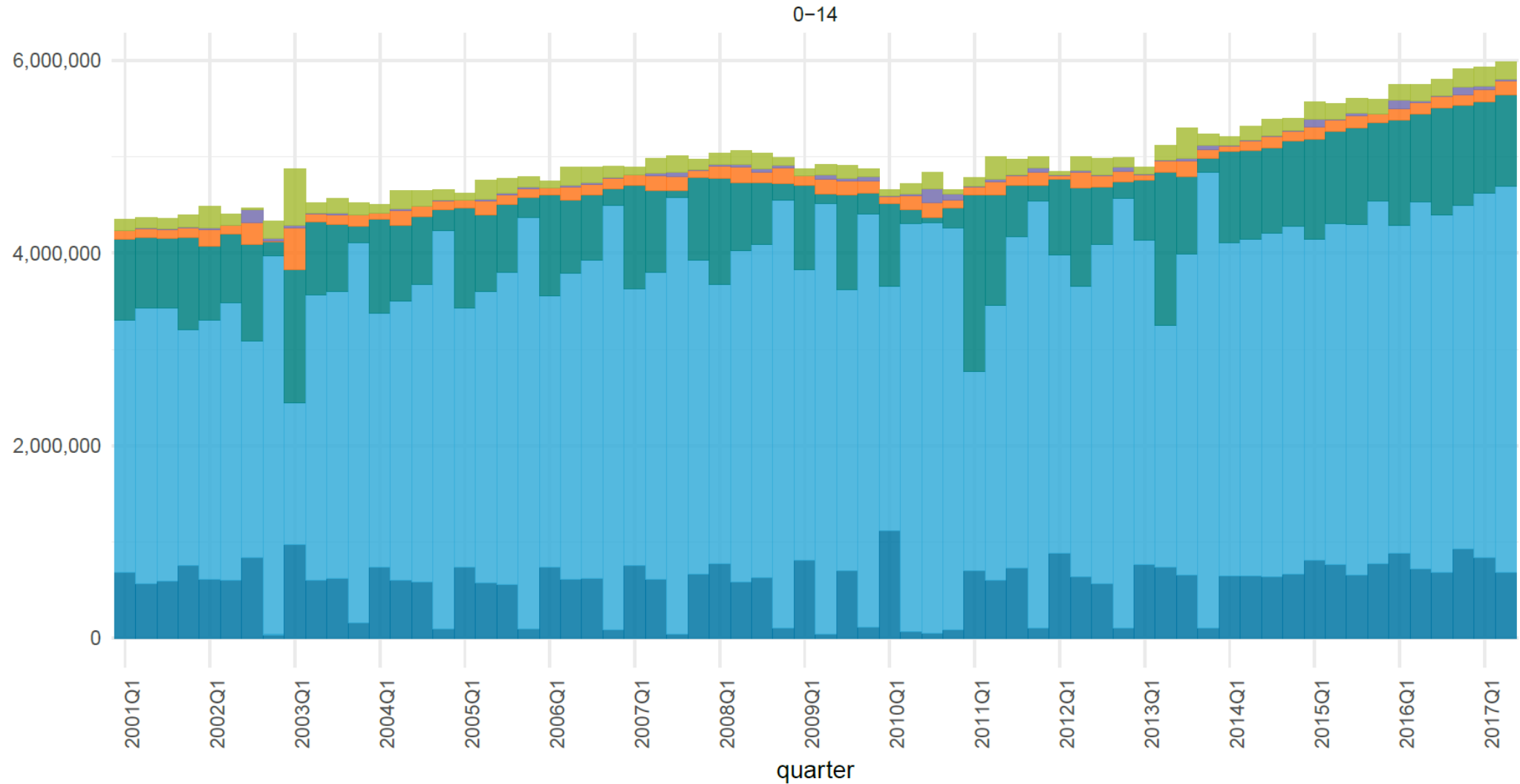
Levels

- Continuers that are growing
- Continuers that are shrinking
- Activators in reference quarter
- Closers about to leave active population
- Activators & Closers simultaneously (Act-Deact)

Total workers in active enterprises, by enterprise transition status



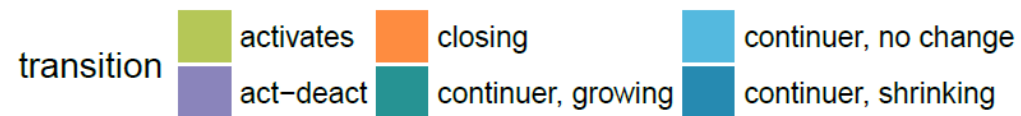
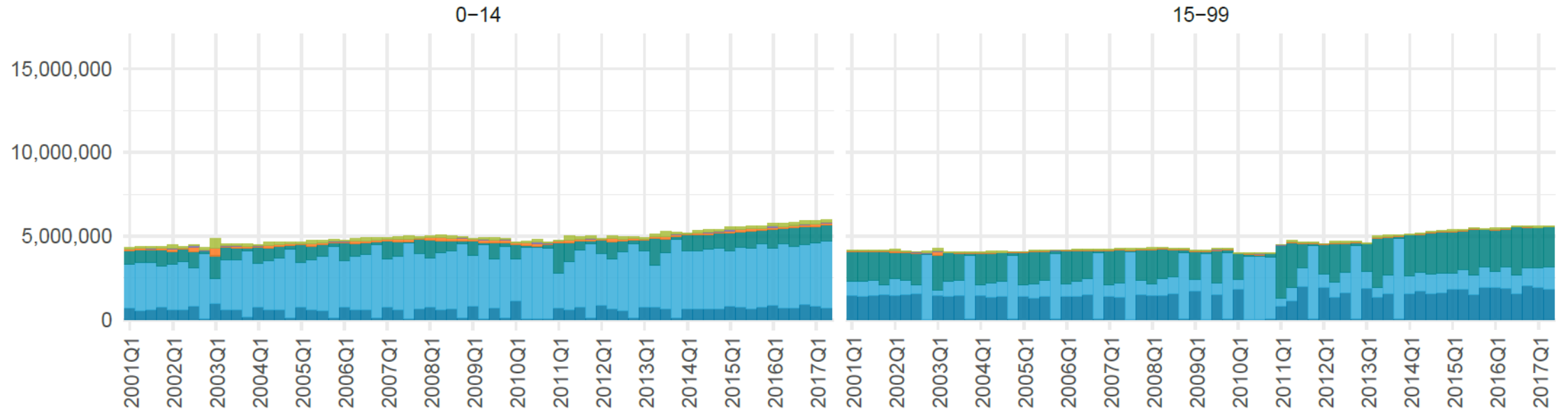
Total workers in active enterprises, by enterprise transition status and sizeband



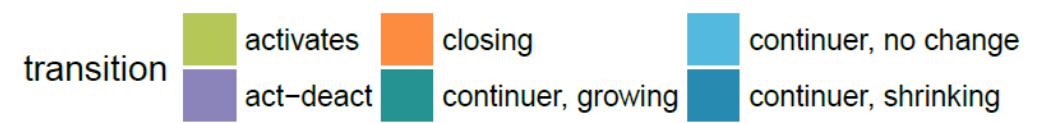
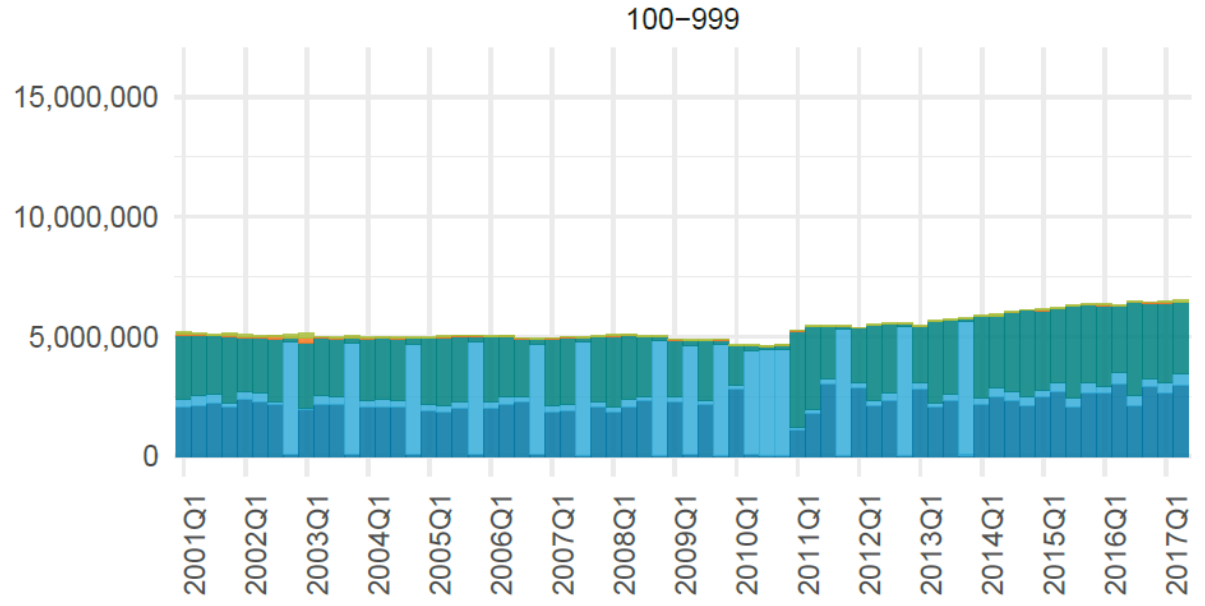
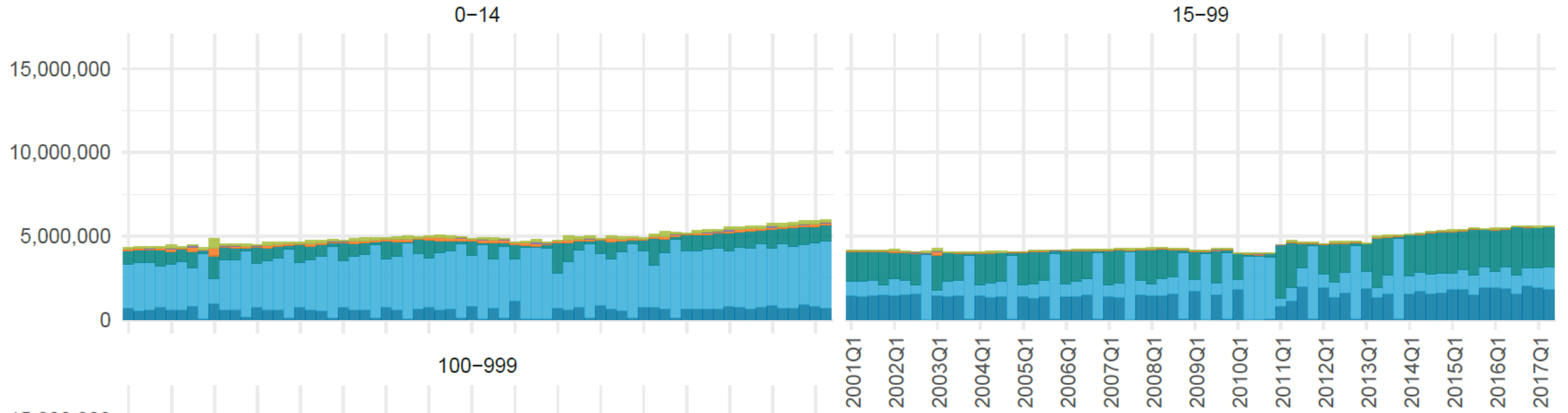
transition

- activates
- act-deact
- closing
- continuer, growing
- continuer, no change
- continuer, shrinking

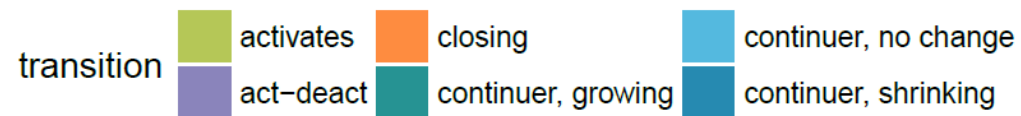
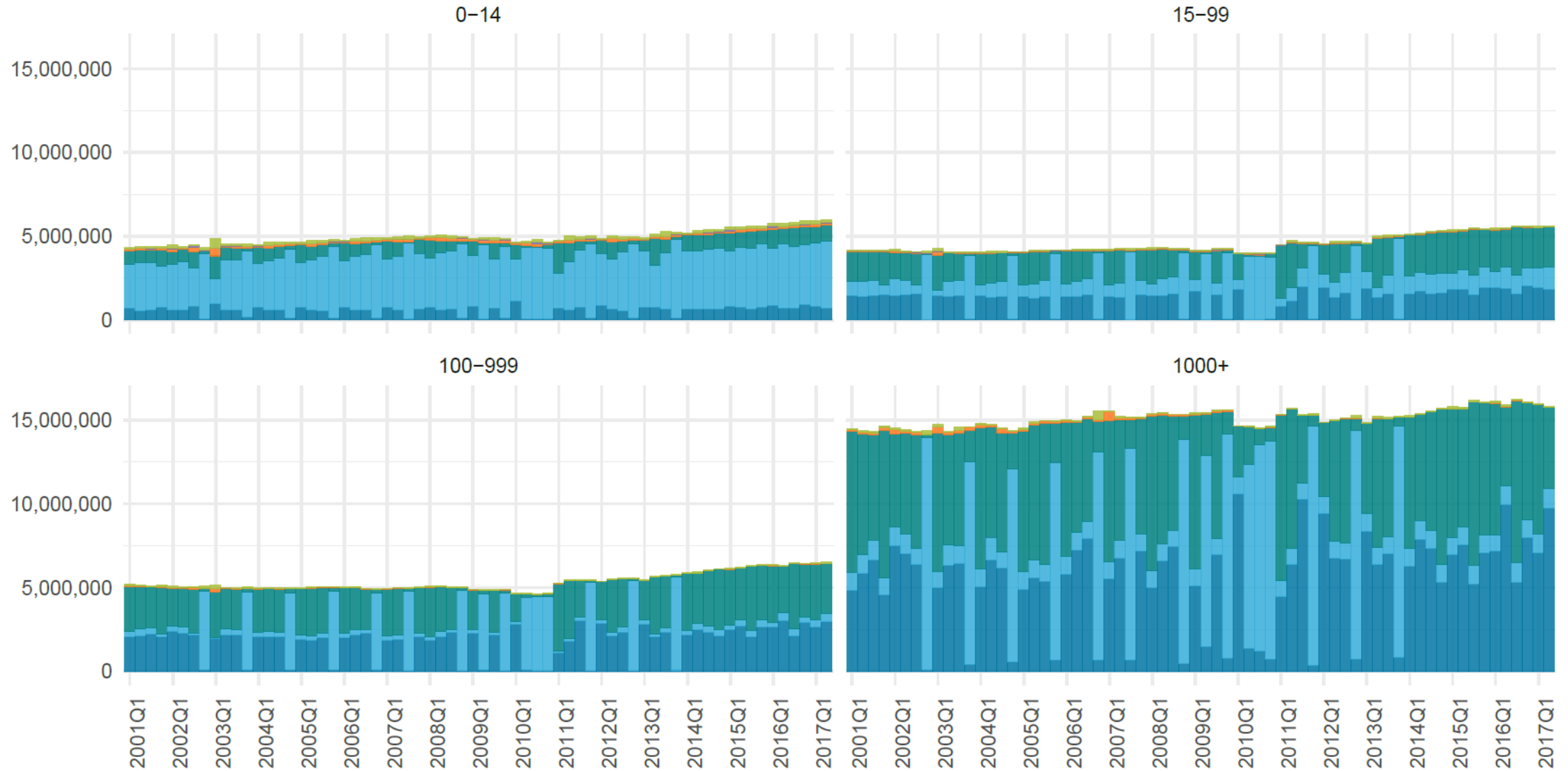
Total workers in active enterprises, by enterprise transition status and sizeband



Total workers in active enterprises, by enterprise transition status and sizeband



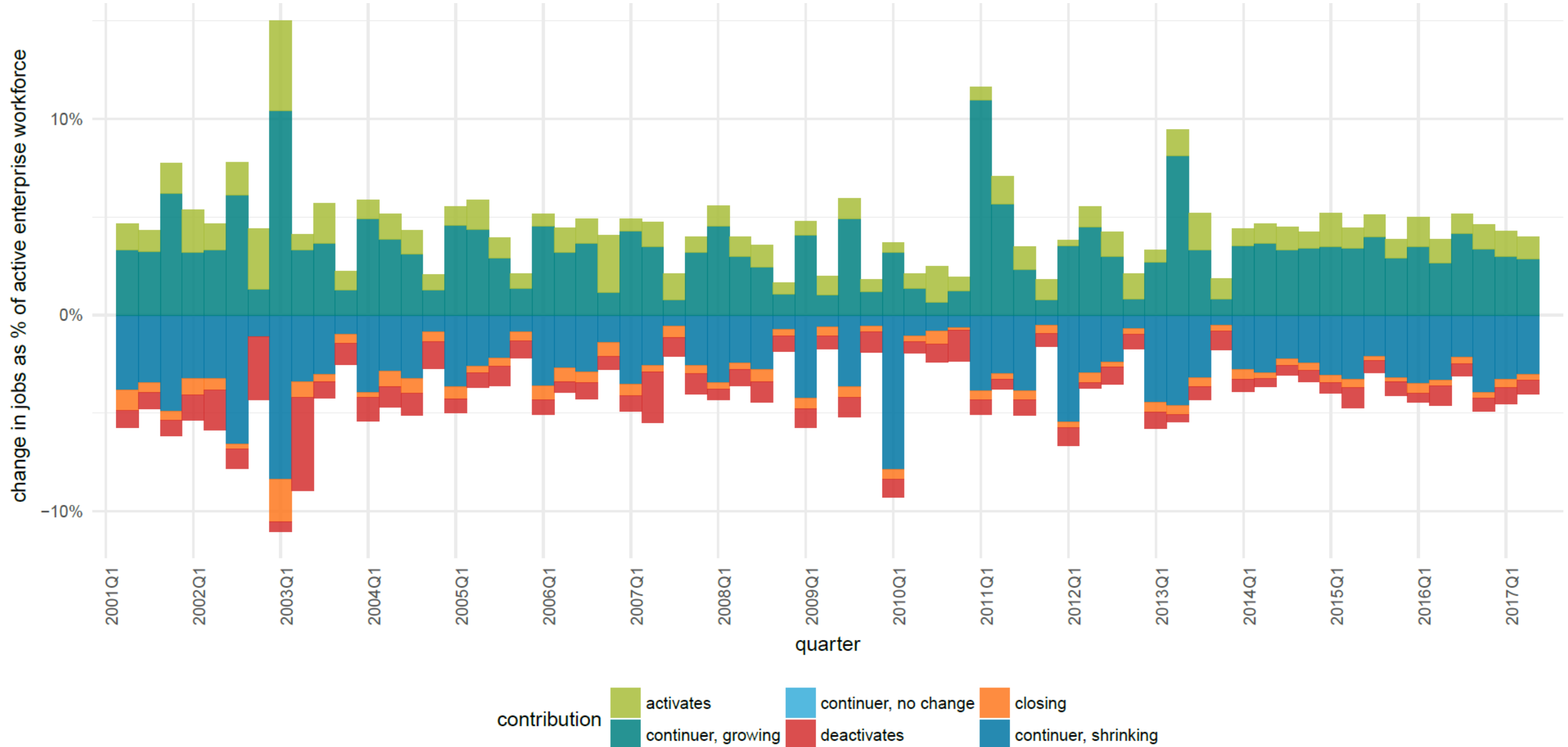
Total workers in active enterprises, by enterprise transition status and sizeband



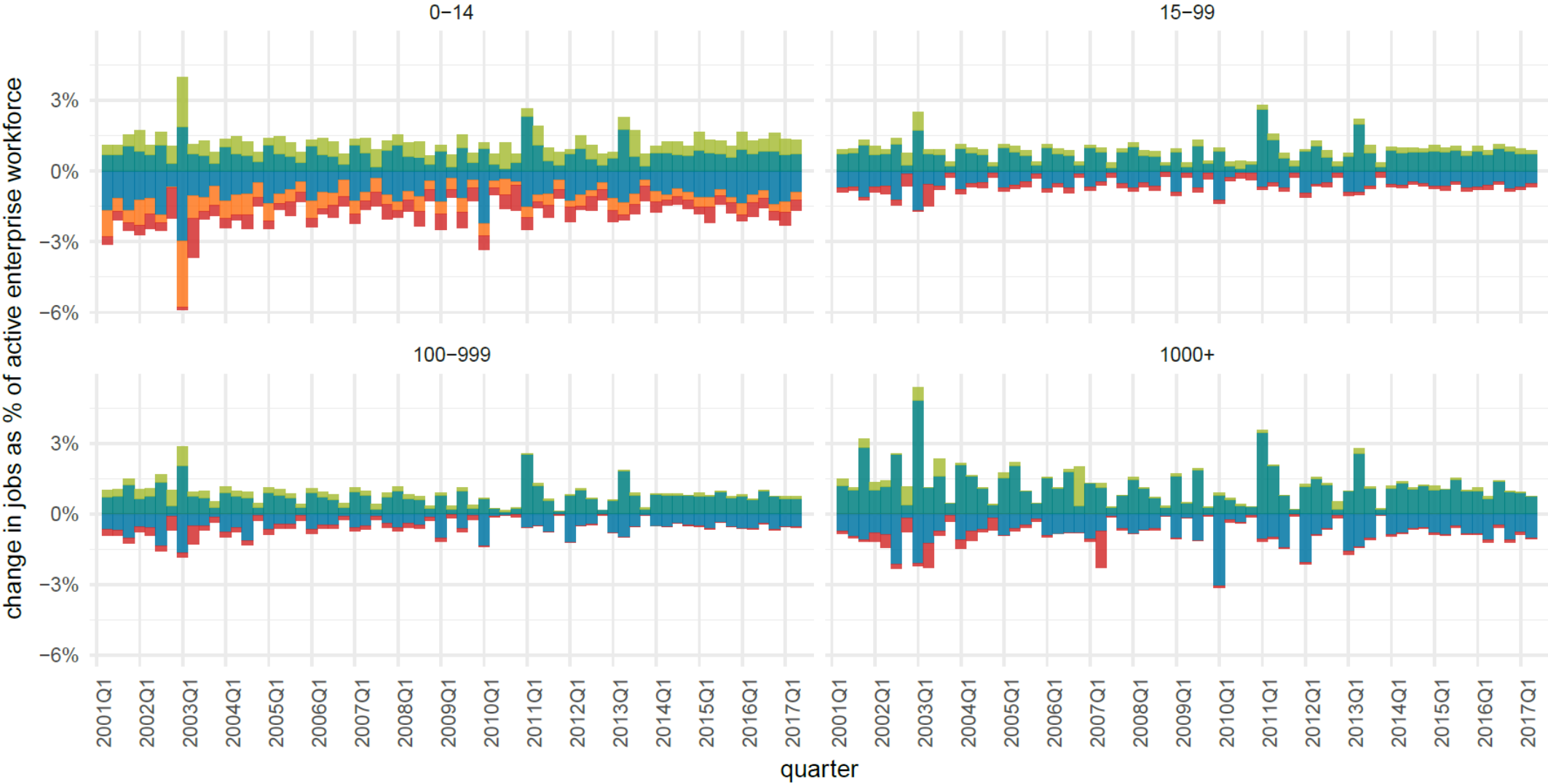
Contributions

- Activations
- Deactivations
- Net contribution from closers (in the period before they deactivate)
- Net contribution from continuers that are growing
- Net contribution from continuers that are shrinking

Total contributions to employment in active enterprises

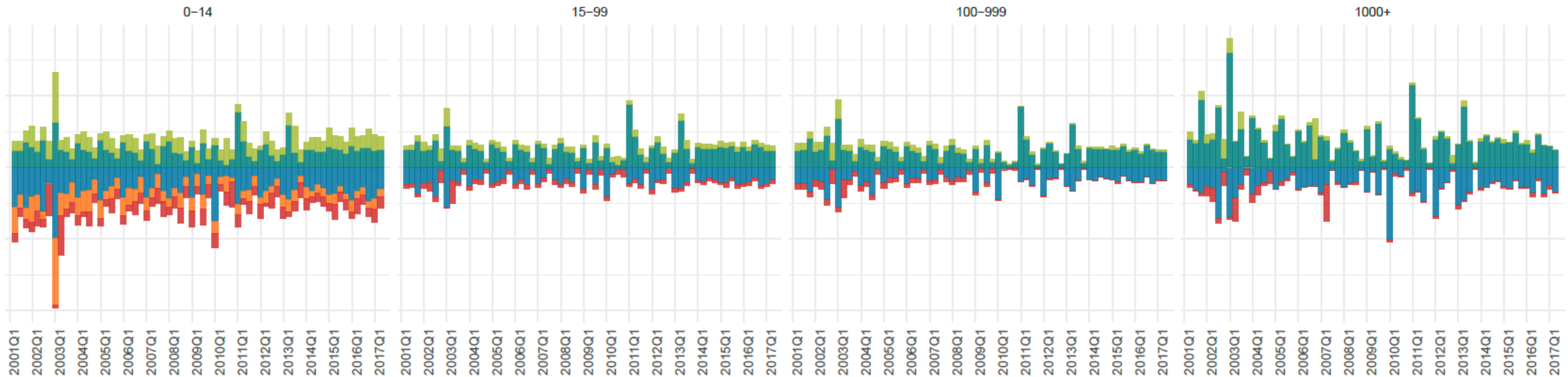


Total contributions to employment in active enterprises by sizeband

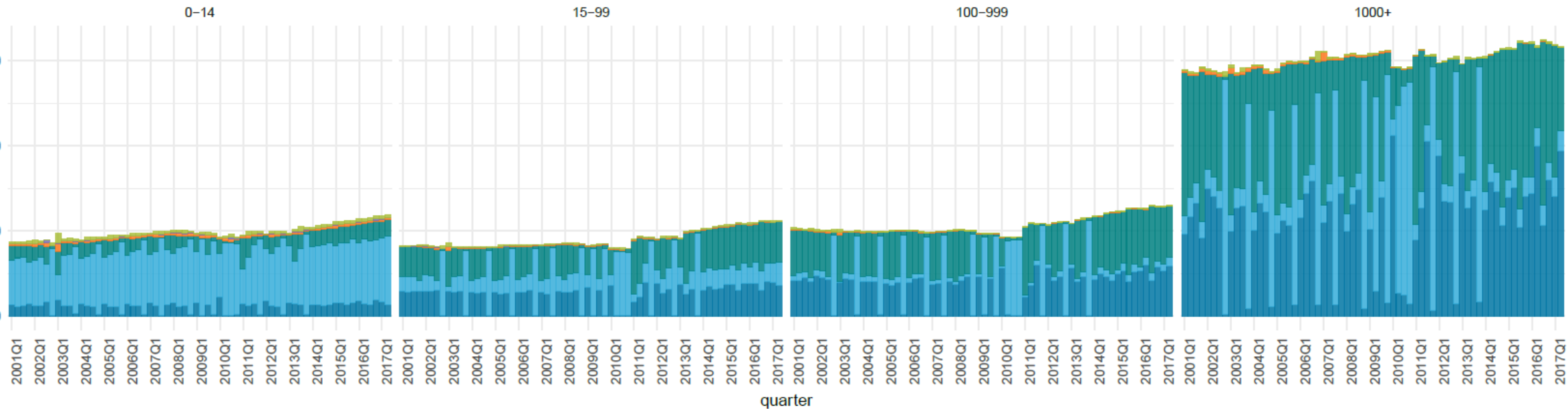


Levels and contributions

change in jobs as % of active enterprise workforce



active enterprise workforce by transition status



transition
activates closing continuer, no change
act-deact continuer, growing continuer, shrinking

Average quarterly contribution to change in total employment of active enterprises

	2001-2007	2008-2009	2011-2017
activations	1.45%	0.89%	1.14%
deactivations	-1.24%	-0.84%	-0.68%
closing firms (net)	-0.61%	-0.45%	-0.36%
continuing and growing firms	3.56%	2.78%	3.31%
continuing and shrinking firms	-3.04%	-2.30%	-2.91%
sum	0.13%	0.07%	0.50%

	2001-2007	2008-2009	2011-2017
activations	1.45%	0.89%	1.14%
deactivations	-1.24%	-0.84%	-0.68%
closing firms (net)	-0.61%	-0.45%	-0.36%
continuing and growing firms (net)	3.56%	2.78%	3.31%
continuing and shrinking firms (net)	-3.04%	-2.30%	-2.91%
sum	0.13%	0.07%	0.50%

Average quarterly contribution to change in total employment of active enterprises

		2001-2007	2008-2009	2011-2017
0-14	activations	0.58%	0.48%	0.56%
0-14	deactivations	-0.48%	-0.51%	-0.42%
0-14	closing firms (net)	-0.71%	-0.45%	-0.36%
0-14	continuing and growing firms (net)	0.75%	0.62%	0.72%
0-14	continuing and shrinking firms (net)	-1.09%	-0.85%	-0.97%
15-99	activations	0.23%	0.18%	0.20%
15-99	deactivations	-0.21%	-0.18%	-0.14%
15-99	closing firms (net)	-0.01%	0.00%	0.00%
15-99	continuing and growing firms (net)	0.73%	0.62%	0.79%
15-99	continuing and shrinking firms (net)	-0.58%	-0.46%	-0.56%
100-999	activations	0.25%	0.10%	0.17%
100-999	deactivations	-0.18%	-0.08%	-0.06%
100-999	closing firms (net)	0.02%	0.00%	0.00%
100-999	continuing and growing firms (net)	0.73%	0.57%	0.74%
100-999	continuing and shrinking firms (net)	-0.56%	-0.42%	-0.53%
1000+	activations	0.40%	0.14%	0.21%
1000+	deactivations	-0.37%	-0.07%	-0.06%
1000+	closing firms (net)	0.10%	0.01%	0.00%
1000+	continuing and growing firms (net)	1.35%	0.96%	1.07%
1000+	continuing and shrinking firms (net)	-0.80%	-0.58%	-0.86%
	sum	0.13%	0.07%	0.50%

Limitations of our analysis

Caveats:

- Enterprises defined as legal entity. We are counting M&A, legal restructuring as job growth/destruction.
- Methods for incorporating PAYE data into IDBR not fully consistent over time
- Do not know gross changes at enterprise level
 - (firm with +100, -50 looks the same as firm with +50)

Industry-level

- “Absolute churn rate”: sum of average absolute contributions in either direction
- Change in dynamism broad-based across industries
- Highest in contrast in the two periods in manufacturing and retail
- ***Caveats:*** Picture varies for different levels of industry aggregation

Average contribution to change in employment

	Activations		Deactivations and closing firms		Continuing firms, growing		Continuing firms, shrinking		Difference in total absolute contribution rates	Workers in 2011 (millions)
	2001-2007	2011-2017	2001-2007	2011-2017	2001-2007	2011-2017	2001-2007	2011-2017		
Retail, food, accommodation	1.7%	1.1%	-1.4%	-0.8%	4.3%	3.6%	-3.8%	-3.1%	-2.6pp	5.0
Manufacturing	1.1%	0.9%	-1.4%	-0.6%	3.0%	2.5%	-3.3%	-2.3%	-2.4pp	2.4
Professional services, technology, media	2.5%	2.0%	-2.3%	-1.2%	3.9%	3.8%	-3.4%	-3.1%	-1.9pp	3.6
Construction	2.4%	2.0%	-1.9%	-1.7%	4.7%	4.0%	-4.0%	-3.4%	-1.8pp	1.3
Distribution, transport, wholesale	1.1%	1.0%	-1.2%	-0.8%	3.2%	2.8%	-2.7%	-2.2%	-1.4pp	2.7
Finance	0.9%	1.1%	-1.2%	-0.9%	2.3%	1.4%	-1.6%	-1.4%	-1.3pp	2.1
Other services	2.5%	1.6%	-1.6%	-1.1%	4.6%	5.0%	-3.8%	-4.5%	-0.4pp	4.2
Government, education, healthcare	0.5%	0.5%	-0.3%	-0.3%	2.0%	2.3%	-1.6%	-2.1%	+0.7pp	8.5

Questions?